UNIVERSITY of WEST FLORIDA

Introducing UWF-ZeekData: Network Datasets Based on the MITRE ATT&CK Framework

Dustin Mink, *Sikha Bagui, & Subhash Bagui*

University of West Florida, USA

Agenda

- Introduction
- Datasets, Cyber Range, and Big Data Platform
- Cybersecurity Classes
- Correlation
- Questions

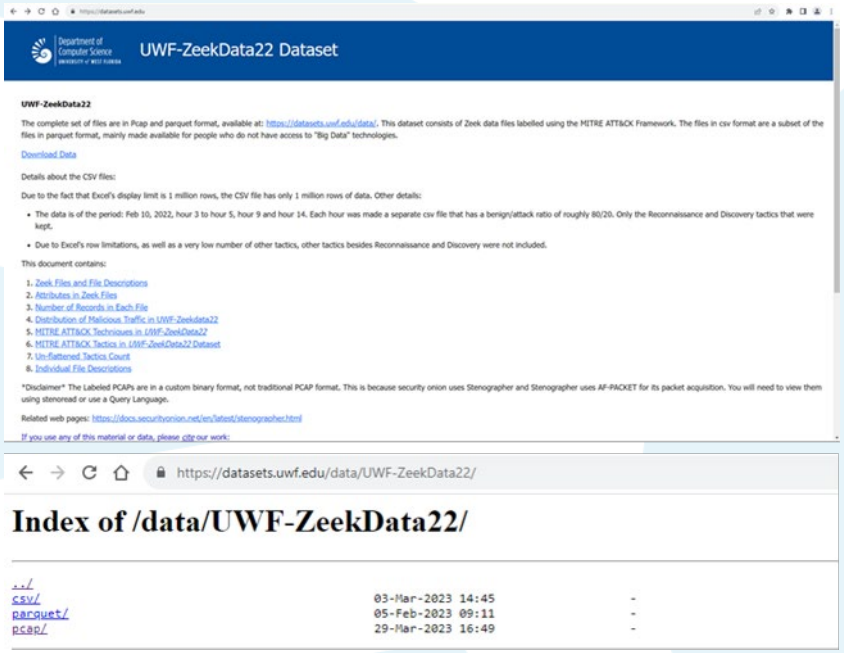
Introduction

- Review major network instruction detection dataset
- UWF-ZeekData22 is the first network instruction detection dataset of Zeek logs labelled with the MITRE ATT&CK framework

UNIVERSITY of WEST FLORIDA

Datasets, Cyber Range, and Big Data Platform

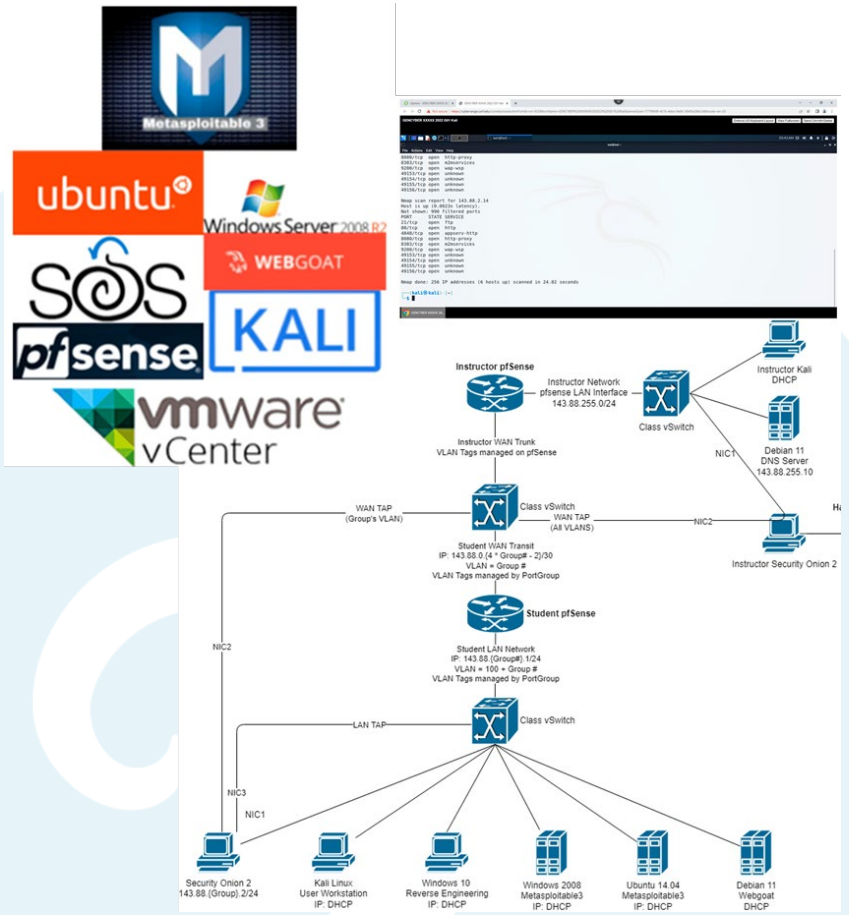
Datasets



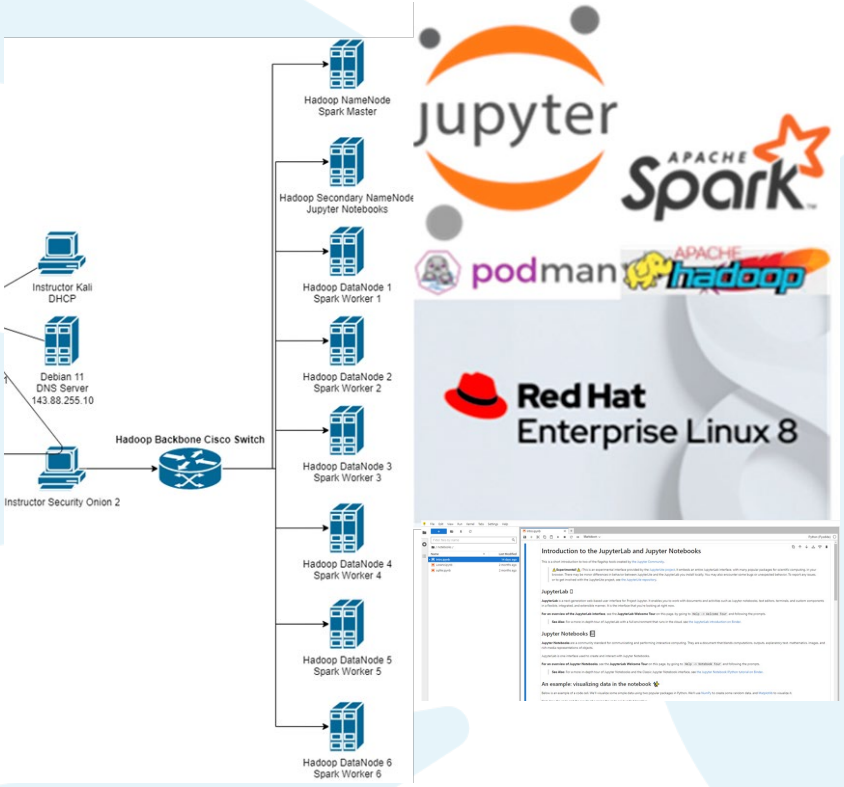
- Publically available web site
- Zeek and MITRE ATT&CK in CSV and Parquet formats
- Raw network traffic in PCAP

UWF Cyber Range

- Supermicro X9DRE-TF+/X9DR7-TF+ (24CPU 128GB RAM 20TB HDD)
- Dell PowerEdge R740 (48CPU 768 GB RAM 6.74TB SSD)
- Dell PowerEdge R750 (128CPU 1TB RAM 12.43TB SSD)
- VMware vCenter
- VMware PowerCLI
- Kali, Security Onion, Pfsense, Metasplotiable 3 (Windows/ Ubuntu), WebGoat
- Ethical Hacking and Penetration Testing and Cyber War Gaming



UWF Big Data Platform



- (×2) Dell PowerEdge R730 (20CPU128GB RAM 4TB HDD)
- (×6) Dell PowerEdge R730xd (20CPU 128GB RAM 48TB HDD)
- Hadoop, Spark, Jupyter Notebooks
- Introduction to Big Data Analytics and Advanced Big Data Analytics

UNIVERSITY of WEST FLORIDA

Cybersecurity Classes

Ethical Hacking and Penetration Testing

- CAE-CD Cybersecurity Program
- Course Description
- Each student has their own Kali VM
- Victims Windows Metasploitable 3, Ubuntu Metasploitable 3, and Ubuntu WebGoat

Cyber War Gaming

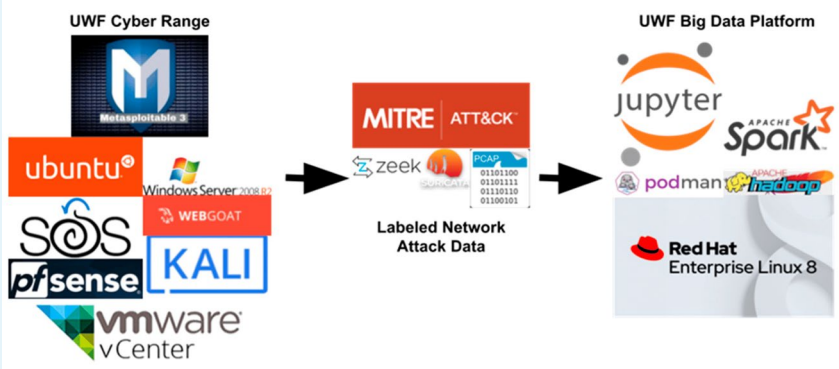
- CAE-CD Cybersecurity Program
- Course Description
- Pairs of Student team up for Capture the Flag
- Kali and Security Onion VMs

UNIVERSITY of WEST FLORIDA

Correlation

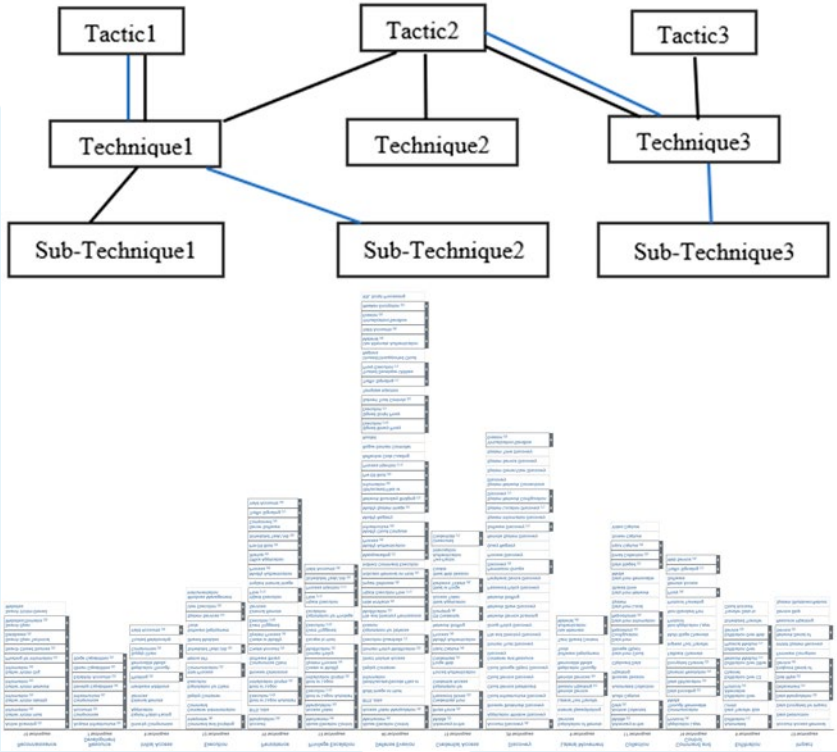
Zeek, PCAP, and Mission Logs

- Instructure Security Onion 2 Collects Zeek and PCAPs
- CronTab runs Bach Script to transfer Zeek and PCAPs to HDFS
- Student enter metadata into Google Form
- End of the semester transfer mission logs to HDFS



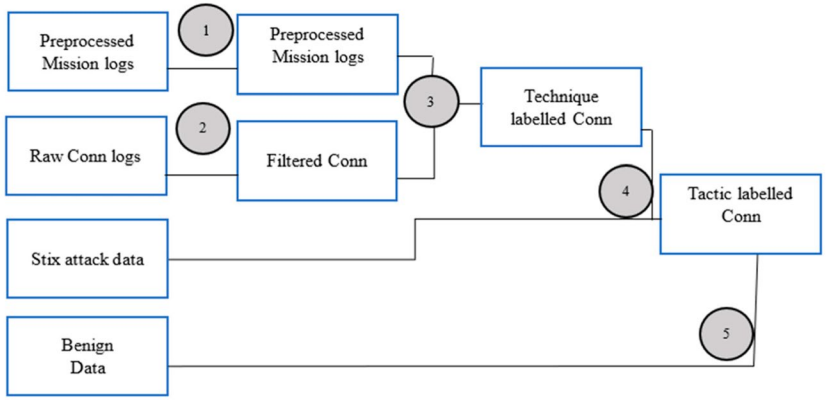
MITRE ATT&CK

- MITRE ATT&CK cybersecurity industry standard
- UWF-ZeekData22, has 14 tactics,191 techniques, and 358 sub-techniques

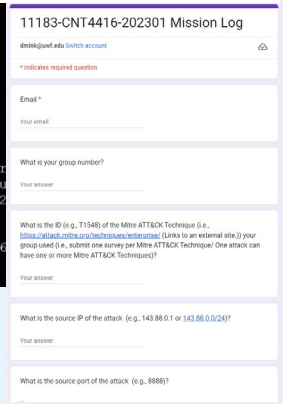


Correlate Zeek and Mission Logs

- Zeek Log
- Mission Log
- Start datetime, stop datetime, src ip, dest ip, src port, and dest port




```
#separator \x09
#set_separator  ,
#empty_field    (empty)
#unset_field    -
#path   conn
#open   2019-08-12-16-26-16
#fields ts      uid     id.orig_h       id.orig_p       id.resp_h
vice    duration        orig_bytes      resp_bytes      conn_state
missed_bytes    history orig_pkts       orig_ip_bytes   resp_pkts
nts
#types  time    string  addr    port    addr    port    enum    str
nt      string  bool    bool    count   string  count   count   cou
1517336042.090842       CBgZ5849tCIsWiMH39      10.55.182.100   142
-       3.000158        0       0       S0      -       -       0
0       (empty)
1517336042.279652       CAcrIotI0pfpNAS13       192.168.88.2    55
dns     0.069982        61      81      SF      -       -       0
109     (empty)
```



UWF-ZeekData22 Data Schema

```
>>> df_conn.printSchema()
root
 |-- resp_pkts: integer (nullable = true)
 |-- mitre_attack: string (nullable = true)
 |-- service: string (nullable = true)
 |-- orig_ip_bytes: integer (nullable = true)
 |-- local_resp: boolean (nullable = true)
 |-- missed_bytes: integer (nullable = true)
 |-- proto: string (nullable = true)
 |-- duration: double (nullable = true)
 |-- conn_state: string (nullable = true)
 |-- dest_ip_zeek: string (nullable = true)
 |-- orig_pkts: integer (nullable = true)
 |-- community_id: string (nullable = true)
 |-- resp_ip_bytes: integer (nullable = true)
 |-- dest_port_zeek: integer (nullable = true)
 |-- orig_bytes: integer (nullable = true)
 |-- local_orig: boolean (nullable = true)
 |-- datetime: timestamp (nullable = true)
 |-- history: string (nullable = true)
 |-- resp_bytes: integer (nullable = true)
 |-- uid: string (nullable = true)
 |-- src_port_zeek: integer (nullable = true)
 |-- ts: double (nullable = true)
 |-- src_ip_zeek: string (nullable = true)
```

- Zeek Connection Log
- MITRE ATT&CK Tactic

Conclusion

- UWF-ZeekData22 is the first network intrusion detection dataset of Zeek logs labelled with the MITRE ATT&CK framework

Acknowledgements

- The Research was supported by 2021 NCAE-C-002: Cyber Research Innovation Grant Program, Grant Number: H98230-21-1-0170

UNIVERSITY of WEST FLORIDA

Questions?

Data available at: https://datasets.uwf.edu